As of February 13, 2017

Q57. **Is TA5 responsible for providing computing hardware for operational evaluations at government sites?**

A57. TA5 proposers should assume that they will be responsible for providing computing hardware to run the Causal Exploration system during government-hosted operational evaluations and during TA5-hosted collaborative experiments, but not during TA4-hosted technology assessments.

As of February 2, 2017

Q56. **Are technical approaches (e.g., deep learning methods) that require training data acceptable?**

A56. Yes, DARPA recognizes that many relevant algorithmic approaches to the problems posed by Causal Exploration will require the use of training data. Proposals incorporating such approaches should clearly describe the nature of required training data and its expected source. Proposals should also specify whether training data are required to: 1) construct and tune algorithms during technology development, 2) initialize/seed the system prior to deployment, or 3) use the system during live operations; and whether new training data is required for each operational environment (OE). Approaches which require OE-specific training data, or that require training data during system operation, must justify the assumption that such training data would be available.

As of January 17, 2017

Q55. **Would a TA2 approach that employs SMEs during Phase 3, off-line and without requiring planner interaction, be in scope (e.g., to incorporate new encoded expert knowledge that was discovered during previous uses of the system)?**

A55: Depending on the specifics of the technical approach, off-line use of SMEs for Phase 3 could be in scope.

Q54. **Can you provide guidance on the total available funds and or/parameters for proposal amounts for the TAs?**

A54: DARPA is not providing a specification of funding allocation.

Q53. **From page 17 of the BAA, "the TA5 performer will provide a processing facility to be used by KO, CMA, and HMI performers for independent development and testing, and to facilitate distribution and exchange of data sets and software prototypes." Does that mean that all development work funded by Causal Exploration must be performed on TA5 storage and compute servers? Also, is it acceptable to propose some equipment purchases under TA1, TA2, or TA3?**

A53: While the TA5 processing facility will be available to the other TAs, it is not required that they use it for component development. Equipment purchases for local development are acceptable but must be justified in the proposal and appropriate to the technical area.

Q52. **Which military message types are TA1 performers to process? Is the TA5 performer (or the Government) to provide a normalization procedure to render the required military message types into a single, standard format?**

A52: DARPA has not identified a specific set of messages, but the intent is that messages representing short text reports may be included. Proposers should state assumptions about pre-processing of messages (e.g. one message per text file).

Q51. **In TA1 results, should we assume that all the causal relationships to be discovered are specific? Or is there a requirement to discover generic causal relationships as well?**

A51: Both general and specific causal relationships should be discovered where there are data in the input sets to support them.

Q50. **Could you identify an example(s) of numeric/databases that would be provided? Are causal relationships to be extracted from the numeric/database data?**

A50: United Nations Statistic Division (UNSD), http://unstats.un.org/unsd/default.htm, the GDELT project, http://www.gdeltproject.org/, and Integrated Crisis Early Warning

System (ICEWS), https://dataverse.harvard.edu/dataverse/icews are example open source databases from which input data may be drawn. TA1 approaches should incorporate numeric/database data as well as textual data in the extraction of causal relationships.

Q49. Could you clarify the statement on page 10: "While programs may consider social media content as a source for knowledge extraction, novel social media analytics are not a primary interest of the program; any such analysis should be included only in the context of a broader capability for extraction of knowledge from open sources?" Could you identify more clearly what is in scope and what is out of scope regarding social media?

A49:   The language in the BAA expresses DARPA's intent.

Q48. To help clarify the boundaries between TA1 and TA2, would research to develop algorithms to address the following topics be considered in-scope for the Causal Model Assembly (TA2) topic?

   a) **Inference of causal relationships not included in TA1 output from knowledge of events that have occurred in the target operational environment**
   b) **Inference of causal relationships not included in TA1 output from statistical data concerning the target operational environment**
   c) **Modification of causal model components from knowledge of events that have occurred in the target operational environment**
   d) **Modification of causal model components from statistical data concerning the target operational environment**
   e) **Determining which existing causal model components are effective for use in modeling incoming data**
   f) **Instantiating causal rules that are independent of the operational environment to create a grounded (i.e., instantiated or localized) causal model**
   g) **Inference of latent factors in a causal model**

A48:   As noted previously, there is not an intention to provide or develop a constructive model of the operational environment and there is no expectation that Causal Model Assembly will have knowledge of events or statistics independent of that provided by TA1. With respect to a) and b), the primary responsibility for inferring causal relationships is expected to reside with the TA1 performers. Identifying additional

causal relationships should not be a major TA2 research focus. With respect to c) and d), modification of causal model components based on events and data provided by TA1 would be in-scope for TA2. The remaining topics could all be considered appropriate in a suitable TA2 proposal.

Q47. **Can TA2 teams expect to receive (from TA1 teams or another source) (1) a set of quantitative variables for possible inclusion in causal models and (2) values for at least some of those variables after each intervention? If not, where do you expect TA2 performers to obtain candidate variables and their values?**

A47:   The BAA describes the expected input sources for the Causal Exploration system. Each TA2 proposer should make clear their expectations for the information to be received from TA1 as a result of TA1's processing of that input. There is not an intention to provide or develop a constructive model of the operational environment that can provide data reflecting the outcomes of arbitrary interventions.

Q46. **The Causal Exploration BAA makes mention of the Athena system as an existing technique requiring considerable modelling expertise. While it's made clear that the effort involved in developing a localized model for Athena is too expensive, are there other deficiencies in the Athena system that make DARPA uninterested in it as an Integrated Causal Model or a substitute for such?**

A46:   Athena embodies many of the desired capabilities and, as the source code is freely available, approaches which propose to incorporate all or part of the existing Athena capabilities would be considered in scope. The current version of Athena does not meet all objectives, including most importantly the ability to provide "backward" causal reasoning (identification of underlying causes). TA2 proposers intending to incorporate any pre-existing models (including Athena) must describe assumptions for TA1 outputs and the resulting TA2 efforts required to formulate a domain-specific configuration for those models.

Q45. On Page 13, the BAA states the requirement that "the ICM must evolve to reflect the latest understanding of the operational environment...". Should this be understood as referring to a temporally grounded causal model into which an instance of an event must be added in order to model what has occurred in the environment, or as an abstract causal model that must be modified to reflect what events are possible in the environment?

A45: The question's description of possible interpretations of the ICM appear dependent on a specific concept for implementation of the ICM. At any point in time the ICM should be fully grounded in the operational environment based on the most recent available information.

Q44. The BAA states, "It is the intent of the program that a common set of interfaces will be developed early in Phase 1 to which all TA1 performers will conform." Do you expect the TA1 consensus output representations agreed early in Phase 1 to include a consensus ontology, or merely a consensus on how to represent an ontology?

A44: The goal is for the TA1 performers to agree to use an existing ontology and a common representation for that ontology early in Phase 1.

Q43: For TA4, do FFRDCs need to provide advance sponsor approval or other special documentation to submit a proposal?

A43: Proposals are not being accepted through the BAA for TA4.

Q42: Is the TA4 award competed through the formal BAA process?

A42: No.

Q41: Should a TA1 PI have a clearance, or is it sufficient for other key personnel to have clearances?

A41: Clearances are not required for TA1 personnel.

Q40: **For the development sets for TA1, will an answer key be provided for a portion of the development sets?**

A40: DARPA is still formulating specific plans for the development data sets, and final plans will be dependent on the proposals selected for award. It is unlikely that an "answer key" will be provided in the sense likely intended by the question, but TA1 proposers should describe assumptions and/or recommendations for the development data sets and this will be incorporated in the plans.

Q39: **Is there an anticipated distribution of funds across the five TAs? Can you divulge that distribution?**

A39: DARPA is not providing a specification of funding allocation.

Q38: **Will DARPA be the contracting office? If not, what office is handling contracts?**

A38: AFRL will be the agent for this effort.

Q37: **How should TA1 cost their involvement in experiments conducted by TA5? Is travel required? Length of travel? Number of trips? Location?**

A37: See page 20 of the BAA. All information is provided except the location of the TA5 performer which cannot be known at proposal time. Proposers should assume a reasonable CONUS location for costing purposes, recognizing that it may be unwise to assume a location leading to a very low travel cost estimate.

Q36: **Do you anticipate TA1-3 performers needing to be "detailed" to the TA5 testing location for Phase II? (e.g., like the CRS for PlanX)**

A36: Continuous on-site presence at the TA5 site not expected. One of DARPA's objectives for TA5 is to minimize the burden of integration and test on the other performers.

Q35: **Just confirming – a performer can prime one technical area and serve as a sub-contractor on another TA?**

A35: Yes.

Q34: **The BAA states that each year of the project should be defined separately, however, the first two phases are 18 months. Can you clarify how to scope the phases in the SOW?**

A34: Tasks should be broken out by phase; cost broken down by task as per cost proposal instructions. An amendment is being issued to the BAA language to reflect this.

Q33: **Many programs have suffered, initially, from a lack of data and agreement on interface and formalisms. Please describe how you envision the various participants will collaboratively address these needs.**

A33: It was the intent that the BAA provide a systematic description of addressing these needs. The Schedule and Milestones section of the BAA describes expectation of collaboration among performers, data sets provided by the government, and by performers.

Q32: **Is learning a new/missing causal relationship (based on observation from data, not direct expert or planner input) within the scope of the program?**

A32: Yes.

Q31: **Is recognition of a flawed or incorrect causal relationship (when directly supported by the knowledge sources) something that should be considered during planning? Or just evaluation over time as evidence accrues to indicate the flaw?**

A31: The proposer should describe how their solution will address this situation.

Q30: **Will the program provide SMEs (e.g., representative users) to collaborate with or should we arrange for such to be part of our team?**

A30: SMEs will be provided by the government for experimentation and evaluation. If the proposer anticipates that SMEs will be necessary for development they should include such support in their proposal.

Q29:   Can you characterize or describe the data provided by the government in TA1?

A29:   See page 21 of the BAA.

Q28:   Is there a corpus of data for development and experimentation?

A28:   See page 21 of the BAA.

Q27:   Approximate division of dollars among the TAs? Number of proposal in each TA? Approximate size of each?

A27:   Multiple awards are anticipated. The level of funding for individual awards made under this solicitation has not been predetermined and will depend on the quality of the proposals received and the availability of funds.

Q26:   Are there specific taxonomies of causal statements that you like? That you want to use?

A26:   This is up to the proposer.

Q25:   Social media, blogs, news reports and intelligence reports often have significant image or video content. Are you interested in including these in the KO/TA1?

A25:   No.

Q24:   Does the BAA language, "must demonstrate an understanding of the operational context" and that TA1 teams "need to map entities and causal relationships into an ontology" mean that tech-focuses TA1 and TA2 teams should include operational design SMEs?

A24:   This is up to the proposer.

Q23: **Do you have any guidance/thoughts on the relative size of TA1 vs. TA2 vs. TA3 awards? Any idea of the number of awards, or dollar amount envisioned for each TA?**

A23: The size and level of funding for individual awards made under this solicitation has not been predetermined and will depend on the quality of the proposals received and the availability of funds.

Q22: **Does each TA1 and TA2 performer need to cover the whole problem domain (models in support of planning) or do you expect each of TA1/TA2 to attack part of the problem? E.g., a TA1 could focus on extracting certain types of data or a TA2 may focus on only certain types of models?**

A22: All proposers must fully address requirements of the TA they are proposing to.

Q21: **There is not a separate "accelerated innovation" TA akin to, e.g., Insight. Would you accept a TA1 bid with some unique components from a smaller firm?**

A21: All bids must fully address requirements of the TA they are proposing to.

Q20: **Since TA1 must make estimates about causality, it seems like it needs a version of an ICM. Therefore, is it correct to assume TA1 and TA2 are linked at a model level, not just through an API. Do you think TA2 will need to understand not just the outputs of TA1 API, but the algorithms themselves?**

A20: As described in the BAA, all proposers should include in their proposals assumptions or expectations about interdependencies with other TA functions. For example, TA1 proposals should include a description of their expectations for the interface between TA1 and TA2 and any interdependences. TA2 proposals should describe assumptions and expectations for both TA1/TA2 and TA2/TA3 interfaces. TA3 proposals should describe assumptions and expectations for TA2/TA3 interfaces. However, it is the intention of DARPA that that those interfaces be defined in such a way as to avoid tight coupling between the functions of the different technical areas, as such tight coupling would be unworkable given multiple performers in each of TA1 and TA2.

Q19: **Is the TA1 performer expect to specify ontologies? If so, who will be responsible for alignment across ontologies?**

A19: It would be preferable that all TA1 performers utilize a common base ontology. It is anticipated that the common agreement on the TA1/TA2 interface will include agreement on ontology. Proposers may suggest existing ontologies that might be used for this purpose.

Q18: **Please clarify level of TA1 involvement in secret experiments, if performers have cleared personnel.**

A18: If a TA1 performer has appropriately cleared personnel, they may participate in testing and experiments with classified data.

Q17: **Are multiple teams expected to be funded under TA3 or is only a single team expected to be funded in TA3?**

A17: Per the BAA, DARPA anticipates making a single award in TA3.

Q16: **What is the anticipated funding level for each TA and each Phase?**

A16: DARPA is not providing a specification of funding allocation.

Q15: **Is there a budget for the program?**

A15: Yes.

Q14: **Is there a suggested budget per TA?**

A14: DARPA is not providing a specification of funding allocation.

Q13: **Will abstracts submitted prior to the due date receive feedback earlier?**

A13:   Abstracts are not necessarily reviewed in the order received and so there is no guarantee that an abstract submitted prior to January 13, 2017, will received feedback earlier than one submitted on the due date. Submit abstracts as early as possible to avoid missing the submission cut-off time and date.

Q12: **Will DARPA make any material from seedling projects available?**

A12:   It is DARPA's assessment that releasing material from seedling projects would not benefit proposers and might cause confusion, as these projects explored many technical paths and concepts which were not incorporated in the Causal Exploration BAA.

Q11: **Should military formalisms such as DIME-FIL and PMESII-PT be considered for modeling?**

A11:   These are not required but proposers are encouraged to consider the space of existing formalisms when formulating their approach.

Q10: **Our organization has a TS facility clearance. Will all classified processing happen at the TA5 facility or can classified data sets be couriered to our facility for processing?**

A10:   Proposers to TA's other than TA5 may propose a plan to conduct classified testing at their facilities and should ensure that their proposal includes a realistic estimate for costs associated with such classified processing including any additional hardware, accreditation, and other costs. DARPA will make a judgment of whether this is preferable for the program. Please note that it should not be taken for granted that a facility cleared for use on one program can be used for classified processing on another program, and even when this is permitted it may take significant administrative effort and months of lead time.

Q9: **Should the TA2 PI have a clearance or just key personnel?**

A9:   Clearance by the PI is not necessarily required. The proposer should consider what personnel may need to be involved in classified testing and in engagements with

operational users in classified settings. It is DARPA's objective that all algorithm and software development will be done at the unclassified level.

Q8: **How will integrated system experimentation and evaluation work with multiple TA1 and TA2 performers?**

A8: There are multiple ways of addressing this issue, and it is DARPA's intent that the formulation of standard interfaces early in the program will maximize flexibility in addressing this situation. It may be desirable in some instances to exercise multiple components in parallel, or it may be desirable for some events to create multiple instances of the system embodying different selections of TA1 and TA2 components. In situations where evaluation is not the primary objective such as experimentation or demonstration, it may be desirable to select a single set of components to use for a given event. The approach taken for a given event will be identified by the TA5 System Integrator in consultation with DARPA, based on the number of performers, the specific technical approaches taken, and the objectives of each event. Proposers to TA5 should consider this issue in formulating their approach to system integration.

Q7: **Our organization has novel or unique data collection capabilities which may be directly applicable to the program. Is DARPA interested in these as part of a TA1 or TA2 approach?**

A7: Data collection per se is not a focus of the program and approaches that are heavily dependent on unique data collection capabilities are not of interest. Proposers wishing to include such capabilities must justify the benefit to the program against the risk that such proprietary or unique data collection capabilities may pose a barrier to transition.

Q6: **Will SECRET level data attributes be different than unclassified?**

A6: DARPA's intent is that unclassified data sets will be representative of all characteristics that may appear in classified data sets. However, proposers should consider and account for the need of minor configuration changes to account for different testing environments without making software modifications.

Q5: **For the HMI is a browser based [thin client] approach OK or does DARPA prefer a heavyweight [thick client] approach?**

A5: DARPA is not specifying a preference for HMI architecture. The program is not expected to provide a user interface that is to be operationally fielded, but rather to demonstrate compelling novel concepts for Human Model Interaction. However, design choices should include a consideration for the feasibility of eventual transition. TA3 proposers should specify the target architecture(s) and justify the rationale for the choice.

Q4: **Is there a preference for industry vs. academia?**

A4: No. DARPA is only concerned that proposed teams incorporate all competencies required to fulfill the requirements for a given technical area.

Q3: **Do you expect TA3 performers to do Human Subject Research (HSR)?**

A3: HSR is not anticipated under this program, but that does not prohibit its inclusion in any proposed solution and HSR usage will be dependent on any proposed approach. Proposals should be clear on whether this is expected to be HSR or not. Proposals which include work judged by the proposer to be HSR should reflect schedule requirements for IRB approval. Bear in mind no DoD/DARPA funding can be used toward HSR until ALL IRB approvals are granted, so plan for only non HSR to occur during the approval timeframe.

Q2: **Can you point to any existing models that might be useful resources available to TA2 proposers?**

A2: The applicability of any given existing model will depend on the approach selected.

Q1: **Will DARPA be receptive to a team proposal that does NOT include a large integrator (i.e., instead has a few small businesses)?**

A1: A proposal which fully and credibly addresses the requirements for a given technical area will be considered responsive regardless of the size of the proposed performers. Any proposal reflecting a team of multiple performers should include a description of

the approach for management and coordination of the team and identify the single entity that would receive any award.